2017-12-09

Trajectories of change in Mediterranean Holocene vegetation through classification of pollen data

Fyfe, R::0000-0002-5676-008X

10.1007/s00334-017-0657-4
Vegetation History and Archaeobotany
Springer Science and Business Media LLC

Uncorrected author's copy. Cite as: Fyfe RM, Woodbridge J and Roberts CN (2017) Trajectories of change in Mediterranean Holocene vegetation through classification of pollen data. *Vegetation History and Archaeobotany* doi:10.1007/s00334-017-0657-4

Trajectories of change in Mediterranean Holocene vegetation through classification of pollen data

Ralph M. Fyfe[1], Jessie Woodbridge and C. Neil Roberts

School of Geography, Earth and Environmental Sciences, University of Plymouth, Plymouth, PL4 8AA, UK

[1] corresponding author: ralph.fyfe@plymouth.ac.uk; +44 1752 585929.

Orchid ID 0000-0002-5676-008X

Abstract

Quantification of vegetation cover from pollen analysis has been a goal of palynologists since the advent of the method in 1916 by the great Lennart von Post. Pollen-based research projects are becoming increasingly ambitious in scale, and the emergence of spatially extensive open-access datasets, advanced methods and computer power has facilitated sub-continental analysis of Holocene pollen data. This paper presents results of one such study, focussing on the Mediterranean basin. Pollen data from 105 fossil sequences have been extracted from the European Pollen database, harmonised by both taxonomy and chronologies, and subjected to a hierarchical agglomerative clustering method to synthesise the dataset into 16 main groupings. A particular focus of analysis was to describe the common transitions from one group to another to understand pathways of Holocene vegetation change in the Mediterranean. Two pollen-based indices of human impact (OJC: Oleaceae, *Juglans*, *Castanea*; API: anthropogenic pollen indicators) have been used to infer the degree of human modification of vegetation within each pollen grouping. Pollen-inferred cluster groups that are interpreted as representing more natural vegetation states show a restricted number of pathways of change. A set of cluster groups were identified that closely resemble anthropogenically-disturbed vegetation, and might be considered anthromes (anthopogenic biomes). These clusters show a very wide set of potential pathways, implying that all potential vegetation communities identified through this analysis have been altered in response to land exploitation and transformation by human societies in combination with

other factors, such as climatic change. Future work to explain these ecosystem pathways will require developing complementary datasets from the social sciences and humanities (archaeology and historical sources), along with synthesis of the climatic records from the region.

Keywords

Pollen, classification, cluster analysis, Mediterranean, Holocene, change analysis

Acknowledgements

This research is funded by the Leverhulme Trust. Earlier versions of this paper were presented at the von Post centennial symposium in Stockholm in November 2016 and the International Pollen Congress in Brazil in 2016, and constructive comments from Thomas Giesecke and John Birks were particularly valuable in shaping the analysis. This study is a contribution to the Past Global Changes (PAGES) project and its working group LandCover6k (http://pastglobalchanges.org/ini/wg/landcover6k/intro), which in turn received support from the US National Science Foundation and the Swiss Academy of Sciences. LandCover6k is coordinated by Marie-José Gaillard (Linnaeus University, Kalmar, Sweden). We thank all contributors to the European Pollen Database. We thank the two reviewers whose comments resulted in a much-improved paper.

Introduction

Pollen analysis allows past vegetation to be described, with the signal mediated through taphonomic and depositional processes. A wide variety of methods have been established to translate pollen data into meaningful vegetation units, including biomisation (Prentice et al. 1996), pseudobiomisation (Fyfe et al. 2010; Woodbridge et al. 2014), modern analogue methods for identifying modern equivalents to fossil assemblages (e.g. Gaillard et al. 1994) and quantifying vegetation cover through model-based transformation (Sugita 2007a, b; Gaillard et al. 2010; Trondman et al. 2015). Understanding the processes that lead to changes in vegetation is a next logical step after palaeo-vegetation classification and description. Much effort is focussed on understanding the role of anthropogenic forcing in transformation of vegetation, something that von Post (1946) largely overlooked, in spite of his engagement with archaeological research programmes.

Humans have transformed natural environments over many millennia, from Neolithic farming to modern agriculture. Williams et al. (2015) have argued that technological development has rapidly transformed ecological structure and dynamics, including vegetation (impacting on species richness, evenness, and biomass), resulting in changes from semi-natural or natural systems to human-modified vegetation. Ellis and Ramankutty (2008) describe the transformation from 'natural' to 'anthropogenic' biomes, using the term 'anthromes'. The development of 'anthromes' likely occurred over many millennia (Ellis et al. 2010; Ellis 2011; Ruddiman et al. 2016), but transformation of vegetation at the global scale is not recognised until recent centuries (Ellis 2011). It is possible to identify anthromes at a regional (sub-continental) scale, particularly in areas with longer histories of complex societies. Understanding the development of (and trajectories towards) anthromes can be facilitated by meta-analysis of pollen data across regions and continents (e.g. Fyfe et al. 2015; Trondman et al. 2015, although neither explicitly describe anthromes). The emergence of such novel ecosystems has been a focus of interest for conservation management, but, as Perring and Ellis (2013) have argued, novel ecosystems are frequently ancient rather than recent developments and ecosystem novelty also depends on when the reference baseline is set (Radeloff et al. 2015). In the Mediterranean, sedentary village life began by 8000 BP and agriculture was established across the majority of the European Mediterranean region by 6000 BP (Roberts 2013). By the 4th millennium BP, complex societies operated within a clearly established 'world' system (Butzer 2005), periodically punctuated by collapses of

higher-order socio-political structures. In addition, fire has long been used as a tool of environmental management, notably to encourage re-growth of new vegetation for livestock grazing and browsing (Vannière et al. 2010).

The divergence of "natural" from modern vegetation in the Mediterranean has previously been explored by Collins et al. (2012), by comparison of fossil assemblages from 6000 BP with recent pollen samples from the same sites, and in the earlier BIOME6k project (Prentice et al. 2000; Roberts et al. 2004). Clear differences are described between 6000 BP as a reference baseline, and modern time periods, most notably via increasing amounts of open ground, and the establishment and spread of disturbed and anthropogenically-modified vegetation. This paper seeks to understand transitions from one vegetation state to another using the example of the Mediterranean. It will assess the extent to which pathways/transitions from one state to another have been predictable and replicated between sites. The methodological basis, which is described in greater detail below, involves "unsupervised" data classification, rather than imposed quantification (c.f. biomisation: Prentice et al., 1996), to produce a taxonomy of pollen samples, grouped by similarity in their assemblages. The paper draws on a large number of fossil pollen sequences from across the Mediterranean basin. Detailed analysis that compares the unsupervised classification of these pollen data to pre-existing classifications (e.g. biomisation) and modern vegetation mappings is described in Woodbridge et al. (in review). Recent work by Felde et al. (2014, 2016) has demonstrated the clear potential of such an approach in the analysis of large pollen datasets, something that even a decade ago would not have been possible without the advent of high-performance computing, and the establishment of clear statistical approaches for the analysis of palaeoecological datasets (Birks et al. 2012).

Materials and methods

Modern and fossil pollen datasets

The approach to classification of samples used here is Ward's hierarchical agglomerative clustering method (Ward 1963; Murtagh and Legendre 2014), using Euclidean distance as the dissimilarity measure between samples. Whilst a large number of alternative methods exist for the classification of palaeobotanical data, careful experimentation has shown that different

methods commonly produce similar results (Felde et al. 2014, 2016). As the purpose here is to assess trajectories of change between classes of pollen assemblages, the simpler Ward's method has been employed over more complex approaches; Woodbridge et al. (in review) discuss some alternative approaches to classification of the same dataset. An advantage of a hierarchical approach to clustering data is also the ability to (dis)aggregate at different levels across the dendrogram (Garcia-Madrid et al. 2014). Analyses were undertaken using the R package VEGAN (Oksanen et al. 2016).

All pollen data used have been extracted from open-access online data stores. Sub-fossil pollen count data was obtained from the European Pollen Database (EPD: May 2016 version: Leydet 2007-2017). The EPD is a community-driven database that archives site metadata, chronological information and raw count data (Fyfe et al. 2009). Pollen sites located in areas with characteristic Mediterranean climate and vegetation have been selected (Fig. 1), using definitions from Ozendaa and Borela (2000) and Izdebski et al. (2015). Figure 1B and 1C show that Mediterranean climate space is well represented in both our modern and fossil pollen site data. Each fossil sample extracted from the EPD was assigned a calibrated age estimate on the basis of established core chronologies (Giesecke et al. 2014; Leydet 2007-2017). The pollen count data from each site has been summed into contiguous 200-year time windows from 11000-10800 to 200- -65 cal BP (with AD1950 as present day). Additionally, 'modern' pollen count data was taken from the European Modern Pollen Database (EMPD: Davis et al. 2013), a repository of surface pollen count data, including date of collection, locational information and additional site metadata. The taxonomy of the EPD and EMPD was harmonised and simplified, taking the EPD as the base. Where possible a detailed level of taxonomic resolution was retained, particularly for key indicator species of disturbance (e.g. ribwort plantain: *Plantago lanceolata*); other taxa were combined at a higher taxonomic level owing to differences in recording between analysts across Europe (e.g. all evergreen oak species and varieties). All taxa below 1% of the pollen sum that occurred in fewer than 50 samples were excluded, resulting in 260 pollen taxa. The harmonisation and 'binning' of sub-fossil and modern samples into 200-year long time windows resulted in a single dataset of 4164 samples, of which 1610 are modern and 2554 are sub-fossil (from 105 sites). The number of clusters was chosen through visual inspection of the dendrogram and developments of measures that described intra-cluster compactness and inter-cluster dissimilarity, on the basis of mean Euclidean distance scores.

A phytosociological approach was used to describe clusters identified within the pollen data. The frequency of occurrence of taxa within each cluster was calculated, and the abundance (pollen percentages) calculated using the median and inter-quartile range. Frequency of occurrence was based on a five-point scale, where V represents a taxon occurring in 81-100% of all samples in a cluster, IV = 61-80%, III = 41-60%, II = 21-40% and I = 0-20%. Frequencies of V and IV indicate constant taxa, III common taxa, II occasional taxa and I scarce taxa. This approach does not explicitly resolve issues of differential productivity of pollen (Broström et al. 2008); the interpretation of the community data requires this to be taken into consideration. Clusters have been given names on the basis of the constant taxa to aid in description of the results. Detailed comparison between these names, modern vegetation communities and pre-existing classification of pollen data (e.g. biomisation: Prentice et al. 1996) is presented in Woodbridge et al. (in review). For each sample two established Mediterranean 'human impact' indices have been calculated: OJC (Oleacaeae, *Juglans*, *Castanea*: Mercuri et al., 2013a) and API (consisting of the human impact indicators *Artemisia*, *Centaurea*, Cichorieae [syn. Lactuceae], *Plantago*, cereals, *Urtica*, *Trifolium*: Mercuri et al., 2013b), and these indices have been summarised for each cluster.

In order to identify within-site pathways of vegetation change across consecutive time windows the frequency of change from each pollen cluster to all other clusters was calculated. These data have been visualised in chord diagrams, constructed in R using the Circlize package (Gu 2014). Plots are circular in character, and use line width to indicate the frequency by which samples change from one cluster grouping to another between adjacent time windows, in similar fashion to that used in the "clock-face" diagrams of Walker (1970). In order to aid visualisation, separate chord plots have been constructed for each cluster, and are organised to indicate change routes towards a cluster (i.e. precursor clusters) and destinations of change from clusters. Self-links (i.e. periods of stasis where a site does not change cluster group across adjacent time windows) are ignored, as are gaps in individual site records (i.e. where there are no dated samples within a time window).

Results

Differentiation of clusters within the pollen dataset

Sixteen groups of pollen samples have been extracted through the clustering analysis following visual inspection of the dendrogram (Figure 2) and experimentation with aggregation of samples at different levels. The dendrogram shows the aggregation of these 16 groups into eight meta-groups, and labelling of the clusters reflects both of these groupings (i.e. meta-group 8 is an aggregation of four lower-level clusters, labelled 8.1, 8.2, 8.3 and 8.4). The phytosociological approach has been used to generate ecological meaning for each grouping, drawing on the constant taxa that are present in each cluster. The results of this are shown on Figure 3. Whilst importance is given to the range of values of each constant taxon (shown as box-and-whisker plots with the interquartile range, median, minimum and maximum values indicated), this is moderated by knowledge of relative pollen productivity, including lower producers (typically the herbaceous taxa: Broström et al. 2008) and high pollen producers such as Pinaceae (e.g. Stedingk et al. 2008).

Cluster 1 is an aggregation of four distinct groups that all appear to indicate open, disturbed or human-modified vegetation as reflected in the OJC and API indices (Table 1; Figure 3). Cluster 1.1 has as its constant taxa Poaceae, Ericaceae, Pinaceae, *Quercus* (evergreen taxa) and Oleaceae, but is the least compact cluster that is produced according to the summary Euclidean distance scores (Table 2). Cluster 1.2 has Oleaceae as a constant and dominant taxon and has the highest OJC score (ranked 1/16), with a smaller number of co-constants than other clusters, but including Poaceae and *Quercus* (evergreen taxa). Cluster 1.3 includes as constants a series of steppe taxa, such as Chenopodiaceae, *Artemisia*, other open ground taxa including Poaceae which results in the highest API score; Pinaceae and *Quercus* (evergreen taxa) are also constant taxa. Cluster 1.4 has a suite of open ground taxa within the constants, including Poaceae, *Plantago lanceolata*, Chenopodiaceae, Caryophyllaceae, Asteraceae (subfamilies Asteroideae and Lactucoides), Ranunculaceae, and is one of only two clusters to include Cerealia-type as a constant taxon. Cluster 1.4 is ranked 5/16 and 3/16 for the human impact indicator groups OJC and API.

Cluster 2 is a distinct group whether 16 or 8 clusters are derived from the dataset, and is both very compact and distinct from other clusters as measured by average dissimilarity scores (Table 2). The constant taxa (frequency class V) that dominate the assemblages in this group are *Quercus* (evergreen taxa) and Poaceae, with Pinaceae, Oleaceae and Cistaceae notable in frequency class IV. This group of samples is thus considered to represent evergreen oak woodland/scrub. Cluster 3 includes within the constant taxa Cyperaceae (dominant),

Pinaceae, Poaceae and *Quercus* (deciduous taxa). This group is harder to interpret owing to the possible ecological meanings of Cyperaceae, but Woodbridge et al. (in review) argue that it represents a combination of lowland wetland and upland pasture. The cluster is ranked 6/16 on the API index on the basis of the median score for samples within the group.

Clusters 4 and 5 have as their dominant constant taxon Pinaceae (Figure 3), and all have Poaceae and *Quercus* (deciduous taxa) as co-dominants. Overall these groupings have very low OJC indices. The groupings are distinctive and separate out even at lower numbers of clusters (Figure 2) and this is further reflected in the low intra-cluster average Euclidean distance scores (Table 2). Cluster 4 has high values of Pinaceae, and most likely represents pine forest. Cluster 5 is an aggregation of two groups. Cluster 5.1 has proportions intermediate between clusters 4 and 5.2 (Figure 3), and cluster 5.2 has both a greater number of constant taxa including open-ground indicators such as Chenopodiaceae, Caryophyllaceae and Ranunculaceae, and higher proportions of open-ground indicators. Cluster 5.2 thus has a higher API index, and is ranked 5/16. Cluster 5.1 is described as pine woods, and 5.2 as pine steppe.

Cluster 6 has as its dominant constant *Quercus* (deciduous taxa), and is an aggregation of two groups. Both groups have a large number of constant taxa (across both frequency classes V and IV), and whilst the constant taxa in frequency class V are virtually identical, the groups are differentiated on the proportions of these taxa, with 6.2 having lower proportions of *Quercus* (deciduous taxa) and higher values across the open ground taxa. Cluster 6.2 is also the only other group (other than 1.4) to include Cerealia-type as a constant, albeit at low proportions, and is ranked 7/16 on the basis of the API index.

Cluster 7 is a distinct grouping (Figure 2), and is described as fir forest, as the constant dominant taxon is *Abies*, and the other constant taxa are also largely woodland types. It has the lowest rankings for both the OJC (15/16) and API (16/16) rankings. Cluster 8 is an aggregation of four of the 16 clusters that are also largely dominated by woodland taxa, and as a group have the lowest OJC and API scores. Cluster 8.3 appears to represent beech woodland, 8.4 mesic forest, 8.1 alder woods and 8.2 non-Pinaceae coniferous forest. Whilst these are compact and distinct clusters on the basis of the mean intra-cluster Euclidean distance dissimilarity scores, the lower inter-cluster dissimilarity scores between the four sub-clusters reflect their greater similarity (Table 2).

The spatial distribution of some common cluster groups for the early (11,000-8,200 BP), mid (8,200-4,200 BP) and late (4,200 BP - present) Holocene are shown in Figure 4. The Holocene subdivisions correspond to those defined by Walker *et al.* (2012). Some of the vegetation clusters have strong regional distribution (Figure 4). Sclerophyll parkland (1.1), for example, is most common in Iberia, and the central Mediterranean region shows the clearest late Holocene increase in parkland/grassland (1.4). Although pine woods (5.1) are widespread in southern France and Iberia in the early Holocene, this cluster group is absent in the eastern Mediterranean at this time, and instead appears in the mid and late Holocene. Deciduous oak woods (6.1) are most common in the eastern Mediterranean and parkland/grassland (1.4) is common here throughout the Holocene, rather than showing a late Holocene rise.

Transitions between clusters

The change (transition) from one cluster group to another has been visualised in chord diagrams (Figure 5). These show, in each case, the precursor groups (blue), and the subsequent direction of change (green) for each cluster. Clusters that have few switches either to or from them (<10) have been excluded from the plots (clusters 1.2, 2, 8.1, 8.2 and 8.3). Values on the plots are proportions of the total number of changes i.e. thicker lines indicate more frequent changes in the direction indicated.

The key observation that can be made from the chord plots is that for the clusters that are indicative of more open vegetation and those with higher OJC and API scores (clusters 1.1, 1.2, 1.3 and 1.4), there are no dominant transition routes from one cluster to another. These clusters have a broad spectrum of pathways both to and from them across the range of sites used here. As an example, transitions to cluster 1.1 (sclerophyllous parkland) come from all other clusters with the exception of clusters 8.1 (alder woods) and 8.3 (beech woods). No single route is dominant although some pathways are more common, including 1.2 (evergreen shrubland) and 1.3 (sclerophyllous steppe/parkland). Onwards transitions for cluster 1.1 go to one of nine (from 15) possible clusters. A similar pattern is observed for cluster 1.4 (parkland/grassland), with only cluster 5.1 (pine woods) not acting as a precursor and many recorded 'destination' clusters. The same broad pattern is observed for cluster 3 (pasture/wetland). Almost all possible pathways to this cluster are recorded (with cluster 1.4 the most frequent, but not the dominant route).

In contrast, clusters that represent closed vegetation (i.e. forest, or natural as inferred from the very low OJC and API scores) have fewer possible precursors. For example, cluster 4 (pine forest) largely emerges from one of four clusters: 5.1 (pine woods), 5.2 (pine steppe), 8.4 (mesic forest) and 1.3 (sclerophyllous steppe/parkland). Cluster 6.1 (deciduous oak woodland) comes from one of eight paths, but has dominant routes that include cluster 6.2 (deciduous oak parkland), 2 (evergreen broad-leaved shrubland) and 8.4 (mesic forest). Cluster 7 (fir forest) only develops from cluster 8.4 (mesic forest) and cluster 8.3 (beech forest). A greater number of possible destination clusters exist for these more closed vegetation groups, with the exception of cluster 6.1 (deciduous oak woods) which largely transitions to cluster 6.2 (deciduous oak parkland).

Discussion

Identifying natural and anthropogenic clusters

The value in using palaeoecology to inform resource and conservation managers in relation to vegetation pathways and change under known past environmental change has been identified by many authors (e.g. Birks 1996; Jackson and Blois 2014; Edwards et al. 2017). The approach employed here has allowed us to gain novel insights into past groupings of pollen samples, and by inference the nature of past vegetation, in a way that has not been attempted previously. The results of the meta-analysis of pollen samples has identified clusters that fall along a spectrum of human impact, from those assemblages that resemble more natural communities, and those that show clear levels of anthropogenic transformation, supported by the human impact indices OJC and API (Mercuri et al. 2013a,b). More natural vegetation communities are those clusters characterised by deciduous (or sclerophyllous) woodland, such as clusters 8.1-8.4 (alder, coniferous, beech and mesic forests), 6.1 (oak woods) and 7 (fir forest). The clusters in group 1 have the highest human impact scores from both woodland cultivation (the OJC index) and pastoral and arable indicators (the API). These communities are thus most likely to represent human-modified vegetation and thus 'anthromes' (cf Ellis 2011), as identified through novel groupings of pollen resulting from the transformation of natural vegetation systems. Measures of ecosystem novelty are not included here, although recent work has begun to demonstrate the timing and scale of emergence of novel communities within Europe (Finsinger et al. 2017). A third broad

grouping is those clusters dominated by pine, although at this stage it is difficult to assess the role of natural disturbance regimes (e.g. fire: Morales-Molino et al. 2013) in the development or fragmentation of these pine groups.

Butzer (2005) describes two important land-use systems that emerge within the Holocene in the Mediterranean: upland pastoral land use and Mediterranean polyculture. Mediterranean polyculture represents a diverse set of land use management approaches in a single system, including grain, animals and orchards (olive, walnut and grapes). According to Grove and Rackham (2003) the scale of patches (fields, or blocks of individual types of land use) within the polyculture system can be small and highly fragmented. Thus, any pollen samples from within a cultural landscape may potentially include a variety of these different agricultural land types, and still retain elements of less-disturbed vegetation. The use of key 'indicator' taxa within the pollen clusters (the OJC and API indices) has allowed these transformed vegetation communities to be identified. The use of two indicators of human impact further supports the inferences: the Spearman's rho reveals a highly significant relationship between the two indices ($rs[16] = 0.81$, $p < 0.000$).

Transitions in Holocene Mediterranean vegetation

The analysis of change between cluster groups can provide useful insights into the development of the Mediterranean vegetation. Differences between natural and anthropogenically-modified vegetation can be observed; for example, more natural clusters (e.g. 4, 6.1, 7, 8.4: Figure 5) typically show only a small number of different precursor groups. Thus pine forest (cluster 4: Figure 5) largely develops from pine woodland (expansion of pine), pine steppe (development of woodland) and steppe parkland (early Holocene expansion of pine). Deciduous oak woods (cluster 6.1: Figure 5) largely develops from oak parkland (woodland expansion and development) and mesic forest (suggesting competition with other species), although in a small number of cases it can replace sclerophyllous vegetation types.

In contrast, those clusters identified as reflecting human-modified vegetation (within the broad cluster groups 1 and 3) have a very wide set of potential pathways of change. The transformation of natural vegetation towards these clusters (Figure 5), is thus not predictable.

This suggests that past societies did not preferentially alter one set of vegetation communities over another (e.g. selective fragmentation of mesic forests over pine steppe), but appear to have exploited and transformed a wide array of natural vegetation types. However, vegetation changes are also influenced by the climatic limits and ecotones of different plant communities. Previous comparisons of Mediterranean vegetation at 6000 BP and the present-day confirm the transformation of much of the landscape (Collins et al. 2012). This included not only an overall reduction in woodland cover (reflected in arboreal pollen percentages) but also the existence, and transformation of, xeric communities between 6000 years ago and present. Collins et al. (2012) used this as evidence that the modern Mediterranean is thus not principally a transformation from a largely mesic mid-Holocene forest vegetation to the largely open and fragmented modern landscape. This is supported by the findings presented here: anthropogenic clusters can emerge from many other vegetation clusters.

A notable feature of the chord diagrams is reciprocity of pathways between clusters (Figure 5). For example, cluster 4 (pine forest) is frequently preceded by cluster 5.1 (pine woods). Destinations from cluster 4 are also frequently cluster 5.1. A similar pattern can be observed for clusters 6.1 and 6.2 (oak forest and oak parkland). This is in part a reflection of the methodological basis of hierarchical clustering. All samples must fall within a discrete cluster (as described below), and inevitably similar samples may fall within different classes if they lie close to the divisions imposed on the data in the method. As vegetation rarely changes abruptly in either space or time, temporal autocorrelation between samples in a site may mean that adjacent assemblages are very similar, but 'flicker' (i.e. short-duration shifts) between similar cluster groupings (e.g. oak parkland and oak forest, which are largely differentiated on the basis of proportions of a similar group of constant taxa: Figure 3). Such flickering has been observed in other classification methods with an arbitrary element (e.g. pseudobiomisation, which uses an affinity score to assign pollen samples to predetermined classes: Fyfe et al. 2010). Whilst this is the most likely explanation for the reciprocity between groups, part of the pattern may also reflect changes in intensity of land use, particularly when clusters move to, and from, clusters with higher or lower implied levels of human modification such as oak parkland (API rank 7: Table 1) to parkland/grassland (API rank 3: Table 1), or linked to periods of more/less frequent landscape burning (e.g. Kaniewski et al. 2008). Butzer (2005) has described cycles of intensification and de-intensification within the Mediterranean, which reflect expansion and contraction of complex societies

within a wider set of world systems. Within this framework of intensification/de-intensification the expected pattern of Mediterranean vegetation would not be a monotonic drive towards increased production and thus increased modification of vegetation. Phases typified by re-expansion of more natural vegetation communities (perhaps on ~600 year time scales, according to Butzer) would be expected. This may reflect movement towards, and from, anthropogenic clusters (i.e. oak woodland to parkland with intensification, and back with de-intensification). Butzer (2005) takes as an example the Lake Lerna pollen record from the southern Argolid region of Greece (Jahns 1993). Cycles of disturbance (with associated weed and cereals) and revertance (re-establishment of more natural vegetation) can be clearly linked to the archaeological record from the region. The implications are that whilst movement of samples between broadly similar clusters may reflect the characteristics of the method, they may also reflect real variations in land use intensity.

The value of classification of pollen samples

The unconstrained organisation, or grouping, of data offers opportunities in the description of past vegetation from pollen data, by allowing the inherent structure in the data to emerge, particularly when a priori groupings in the data cannot be established (Felde et al. 2016). Although previous work has also undertaken inductive data exploration such as that described in this paper through classification (notably Huntley 1990) it remains an under-utilised approach (Felde et al. 2014). The approach assumes that distinct groupings exist with the dataset under study. Whilst there are in general clear and distinct vegetation communities (and thus pollen assemblages) both in the present and the past Mediterranean landscape (e.g. pine forests vs deciduous oak woodland) the gradual nature of vegetation change in both space and time means that edges between groups may be 'blurred' (e.g. ecotones between major vegetation communities). Further, pollen records integrate the signal of all vegetation communities within the source area of the pollen. In highly heterogeneous landscapes this may present particular challenges where pollen may be sourced from both highly modified and more natural vegetation communities. In spite of these caveats, it has been possible through this approach to recognise distinct Mediterranean vegetation types as reflected in sets of distinctive pollen assemblages.

The results shown in the chord plots (Figure 5) are a first step in describing general patterns of vegetation development across the Mediterranean basin, focusing on transitions from one state to another. Future efforts might include comparison of Holocene pathways of change with those from previous interglacials, which must reflect natural vegetation communities and pathways of change (such as those described by Tzedakis 2007), to further elucidate human impact on Holocene vegetation. Changes driven by extrinsic factors (e.g. shifts in Mediterranean climate) might be expected to produce largely temporally consistent shifts in vegetation, whilst intrinsic drivers (including site-specific human impacts) may produce shifts that are more localised in both time and space. Clearly understanding such site-based patterns demands an understanding of both regional differences in climate (e.g. Labuhn et al. (in press) for the first millennium AD), and a detailed understanding of changing population dynamics and land use systems from the archaeological and historical record (Perring and Ellis 2013).

The unsupervised classification approach also has clear limitations, not least that all pollen sites are treated equally. There has been no attempt to account for inter-site taphonomic differences, such as scale of vegetation represented between small and large lakes, or depositional biases between lakes and mires. Woodbridge et al. (2014) showed that whilst such differences can produce noise in the analysis of large datasets, generally they do not obscure the resulting signal that is obtained. Similarly, differential pollen production is not accounted for (e.g. Fyfe et al. 2013; Marquer et al. 2014; Trondman et al. 2015). Placing equal emphasis on frequency of occurrence of taxa and relative proportions of taxa, is an attempt to alleviate this problem. Despite the limitations, there is still considerable value in a data driven approach that is not restricted by the availability or applicability of PPEs and allows major patterns in pollen datasets to be identified.

Conclusions

Ward's hierarchical clustering method has provided a simple approach to the division of a large harmonised dataset from 105 sub-fossil pollen sequences and 1610 surface pollen samples in the Mediterranean extracted from the EPD and EMPD. The division of the dataset into sixteen groups has allowed the identification of clusters that are taken to represent more natural vegetation communities (mesic, coniferous and sclerophyllous forest types), and more

421 open land cover types associated with human impact, supported by summary human impact
422 indicators. Some of these anthropogenic land cover types could be taken to represent
423 anthromes, that is, anthropogenically-modified biomes.

424 Clusters that are taken to represent more natural vegetation assemblages have limited
425 pathways of change between states. In contrast, those with a higher level of human
426 modification (as reflected in the OJC and API rankings) show numerous pathways to their
427 development. This implies widespread fragmentation of all community types as a
428 consequence of human impact, rather than selective transformation of particular types of
429 vegetation. Movement of sites from one cluster to a similar group and back may partly
430 represent an artefact of the method, but it is possible that some of these changes represent
431 cycles of intensification/de-intensification of land use, a phenomenon observed in the
432 archaeological record since prehistoric times.

433 Much more work is required to understand in detail the processes behind the transformation
434 and change in the Mediterranean pollen data represented here and this is a non-trivial task. It
435 will necessitate a comprehensive understanding of human land pressure around each site
436 including land use systems and population dynamics (drawn from archaeological and
437 historical sources), and the compilation of spatially-explicit climatic records. The ability to
438 describe vegetation at the sub-continental scale using approaches such as that presented here
439 is clear, and a next step is to develop the independent (i.e. non-pollen) datasets that can help
440 understand the patterns that emerge.

References

Birks HJB (1996) Contributions of Quaternary palaeoecology to nature conservation. Journal of Vegetation Science 7:89-98

Birks HJB, Lotter AF, Juggins S, Smol JP (2012) Tracking Environmental Change Using Lake Sediments, Volume 5: Data Handling and Numerical Techniques. Springer, Dordrecht

Broström A, Nielsen AB, Gaillard M-J, Hjelle K, Mazier F, Binney H, Bunting M-J, Fyfe RM, Meltsov V, Poska A, Räsänen S, Soepboer W, Stedingk H, Suutari H. Sugita, S (2008) Pollen productivity estimates – the key to landscape reconstructions. Vegetation History and Archaeobotany 17:461-478

Butzer KW (2005) Environmental history in the Mediterranean world: cross-disciplinary investigation of cause-and-effect for degradation and soil erosion. Journal of Archaeological Science 32:1773-1800

Collins PM, Davis BAS, Kaplan JO (2012) The mid-Holocene vegetation of the Mediterranean region and southern Europe, and comparison with the present day. Journal of Biogeography 39:1848-1861

Davis BAS, Zanon M, Collins P, Mauri A, Bakker J, Barboni D, Barthelmes A, Beaudouin C, Birks, HJB, Bjune AE, Bozilova E, Bradshaw RHW, Brayshay BA, Brewer S, Brugiapaglia E, Bunting J, Conner SE, de Beaulieu J-L, Edwards KJ, Ejarque A, Fall P, Florenzano A, Fyfe R, Galop D, Giardini M, Giesecke T, Grant MJ, Guiot J, Jahns S, Jankovská V, Juggins S, Karmann M, Karpinska-Kolaczek M, Kolaczek P, Kuhl N, Kunes P, Lapteva EG, Leroy SAG, Leydet M, Lopez Saez JA, Masi A, Matthias I, Mazier F, Meltsov V, Mercuri AM, Miras Y, Mitchell FJG, Morris JL, Naughton F, Nielsen AB, Novenko E, Odgaard B, Ortu E, Overballe-Petersen MV, Pardoe HS, Peglar SM, Pidek I, Sadori L, Seppa H, Severova E, Shaw H, Swieta-Musznicka J, Theuerkauf M, Tonkov S, Veski S, van der Knaap WO, van Leeuwen J, Woodbridge J, Zimny M, Kaplan JO (2013) The European Modern Pollen Database (EMPD) project. Vegetation History and Archaeobotany 22:521-530

Edwards KJ, Fyfe RM, Jackson ST (2017) The first 100 years of pollen analysis. *Nature Plants* 3. doi:10.1038/nplants.2017.1

Ellis EC (2011) Anthropogenic transformation of the terrestrial biosphere. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 369:1010-1035

Ellis EC, Klein Goldewijk K, Siebert S, Lightman D, Ramankutty N (2010) Anthropogenic transformation of the biomes, 1700 to 2000. Global Ecology and Biogeography 19:589-606

Ellis EC, Ramankutty N (2008) Putting people in the map: anthropogenic biomes of the world. Frontiers in Ecology and the Environment 6:439–447

Felde VA, Bjune AE, Grytnes JA, Birks HJB (2014) A comparison of novel and traditional numerical methods for the analysis of modern pollen assemblages from major vegetation–landform types. Review of Palaeobotany and Palynology 210:22-36

Felde VA, Hooghiemstra H, Torres-Torres V, Birks JBH (2016) Detecting patterns of change in a long pollen-stratigraphic sequence from Funza, Colombia – a comparison of new and traditional numerical approaches. Review of Palaeobotany and Palynology 234:94-109

Finsinger W, Giesecke T, Brewer S, Leydet M (2017) Emergence patterns of novelty in European vegetation assemblages over the past 15 000 years. Ecology Letters 20:336-346

Fyfe RM, de Beaulieu J-L, Binney H, Bradshaw RHW, Brewer S, Le Flao A, Finsinger W, Gaillard M-J, Giesecke T, Gil-Romera G, Grimm EC, Huntley B, Kunes P, Kuhl N, Leydet M, Lotter AF, Tarasov PE and Tonkov S (2009) The European Pollen Database: past efforts and current activities. Vegetation History and Archaeobotany 18:417-424

Fyfe RM, Roberts CN, Woodbridge J (2010) A pollen-based pseudo-biomisation approach to anthropogenic land cover change. The Holocene 20:1165-1171

Fyfe RM, Twiddle C, Sugita S, Gaillard M-J, Barratt P, Caseldine CJ, Dodson J, Edwards KJ, Farrell M, Froyd C, Grant MJ, Huckerby E, Innes JB, Shaw H, Waller M (2013) The Holocene vegetation cover of Britain and Ireland: overcoming problems of scale and discerning patterns of openness. Quaternary Science Reviews 73:132-148

Fyfe RM, Woodbridge J, Roberts CN (2015) From forest to farmland: pollen-inferred land cover change across Europe using the pseudobiomization approach. Global Change Biology 21:1197-1212

Gaillard MJ, Birks HJB, Emanuelsson U, Karlsson S, Lagerås P, Olausson D (1994) Application of modern pollen/land-use relationships to the interpretation of pollen diagrams—reconstructions of land-use history in south Sweden, 3000-0 BP. Review of Palaeobotany and Palynology 82:47-73

Gaillard M-J, Sugita S, Mazier F, Trondman, A-K, Broström A, Hickler T, Kaplan JO, Kjellström E, Kuneš P, Lemmen C, Olofsson J, Smith B, Strandberg G, Fyfe R, Nielsen AB, Alenius T, Balakauskas L, Barnekow L, Birks HJB, Bjune A, Björkman L, Giesecke T, Hjelle K, Kalnina L, Kangur M, van der Knaap WO, Koff T, Lagerås P, Latalowa M, Leydet M, Lechterbeck J, Lindbladh M, Odgaard B, Peglar S, Segerström U, von Stedingk H, Seppä H (2010) Holocene land cover reconstructions for studies on land cover-climate feedbacks. Climate of the Past 6:483-499

García-Madrid AS, Molina JA, Cantó P (2014) Classification of habitats highlights priorities for conservation policies: The case of Spanish Mediterranean tall humid herb grasslands. Journal for Nature Conservation 22:142-156.

Giesecke T, Davis B, Brewer B, Finsinger W, Wolters S, Blaauw M, de Beaulieu J-L, Fyfe RM, Gaillard M-J, Gil-Romera G, van der Knaap WO, Kuneš P, Kühl N, van Leeuwen JFN, Leydet M, Lotter AF, Semmler M, Bradshaw RHW (2014) Towards mapping the late Quaternary vegetation change of Europe. Vegetation History and Archaeobotany 23:75-86

Grove AT, Rackham, O (2003) The nature of Mediterranean Europe: an ecological history. Yale University Press, Yale

Gu Z (2014) Circlize implements and enhances circular visualisation in R. Bioinformatics 30:2811-2812

Hijmans RJ, Cameron SE, Parra JL, Jones PJ, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology **25**:1965-1978

Huntley B (1990) European vegetation history: Palaeovegetation maps from pollen data-13000 yr BP to present. Journal of Quaternary Science 5:103-122.

Izdebski A, Holmgren K, Weiberg E, Stocker SR, Büntgen U, Florenzano A, Gogou A, Leroy AAG, Luterbacher J, Martrat B, Masi A, Mercuri AM, Montagna P, Sadori L, Schneider A, Sicre MA, Triantaphyllou M, Xoplaki E (2015) Realising consilience:

How better communication between archaeologists, historians and natural scientists can transform the study of past climate change in the Mediterranean. Quaternary Science Reviews 136:5–22.

Jackson ST, Blois JL (2015) Community ecology in a changing environment: Perspectives from the Quaternary. Proceedings of the National Academy of Sciences 112:4915-4921

Jahns S (1993) On the Holocene vegetation of the Argive Plain (Peloponnese, southern Greece). Vegetation History and Archaeobotany 2:187-203.

Kaniewski D, Paulissen E, De Laet V, Waelkens M (2008) Late Holocene fire impact and post-fire regeneration from the Bereket basin, Taurus Mountains, southwest Turkey. Quaternary Research 70:228-239.

Labuhn I, Finné M., Izdebski A, Roberts N, Woodbridge J (in press) Climatic changes and their impacts in the Mediterranean during the first millennium CE. *Late Antique Archaeology*

Leydet M (2007-2017) The European Pollen Database. http://www.europeanpollendatabase.net/. Accessed 26 May 2016

Marquer L, Gaillard M-J, Sugita S, Trondman A-K, Mazier F, Nielsen AB, Fyfe RM, Odgaard B, Alenius T, Birks HJB, Bjune AE, Christiansen J, Dodson J, Edwards KJ, Giesecke T, Herzschuh U, Kangur M, Lorenz S, Poska A, Schult M, Seppä H (2014) Holocene changes in vegetation composition in northern Europe: why pollen-based quantitative reconstructions matter? Quaternary Science Reviews 90:199-216

Mercuri AM, Bandini Mazzanti M, Florenzano A, Montecchi MC, Rattighieri E (2013a) *Olea, Juglans* and *Castanea*: the OJC group as pollen evidence of the development of human-induced environments in the Italian peninsula. Quaternary International 303:24-42

Mercuri AM, Mazzanti M, Florenzano A, Montecchi MC, Rattighieri E, Torri P (2013) Anthropogenic pollen indicators (API) from archaeological sites as local evidence of human-induced environments in the Italian peninsula. Annali Di Botanica 3:143-153

Morales-Molino C, García-Antón M, Postigo-Mijarra JM, Morla C (2013) Holocene vegetation, fire and climate interactions on the westernmost fringe of the Mediterranean Basin. Quaternary Science Reviews 59:5-17

563 Murtagh F, Legendre P (2014) Ward's hierarchical agglomerative clustering method: which
564 algorithms implement Ward's criterion? Journal of Classification 31:274-295

565 Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara
566 RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2016) vegan:
567 Community Ecology Package. R package version 2.4-2. https://CRAN.R-
568 project.org/package=vegan

569 Ozendaa P, Borela JL (2000) An ecological map of Europe: why and how? Ecology
570 323:983–994

571 Perring MP, Ellis EC (2013) The extent of novel ecosystems: long in time and broad in space.
572 In: Hobbs J, Higgs ES, Hall CM (eds) Novel ecosystems: Intervening in the new
573 ecological world order, Wiley-Blackwell, Chichester, pp 66-80

574 Prentice C, Guiot J, Huntley B, Jolly D, Cheddadi R (1996) Reconstructing biomes from
575 palaeoecological data: a general method and its application to European pollen data at 0
576 and 6 ka. Climate Dynamics 12:185-194

577 Prentice IC, Jolly D, BIOME 6000 Participants (2000) Mid-Holocene and glacial-maximum
578 vegetation geography of the northern continents and Africa. Journal of Biogeography
579 27:507–519

580 Radeloff VC, Williams JW, Bateman BL, Burke KD, Carter SK, Childress ES, Cromwell KJ,
581 Gratton C, Hasley AO, Kraemer BM (2015) The rise of novelty in ecosystems.
582 Ecological Applications 25:2051-2068

583 Roberts N (2013) The Holocene: an environmental history. John Wiley and Sons, London

584 Roberts N, Stevenson T, Davis B, Cheddadi R, Brewster S, Rosen A (2004) Holocene
585 climate, environment and cultural change in the circum-Mediterranean region. In:
586 Batterby RW, Gasse F, Stickley CE (eds) Past climate variability through Europe and
587 Africa. Springer, Netherlands, pp 343-362

588 Ruddiman W, Fuller D, Kaplan J, Ellis E, Kutzbach J, Vavrus S, Tzedakis P, Roberts N, Fyfe
589 R, He F, Lemmen C, Woodbridge J (2016) Late Holocene climate: natural or
590 anthropogenic? Review of Geophysics 54:93-118

591 Stedingk H, Fyfe RM, Allard A (2008) Pollen productivity estimates for the reconstruction of
592 past vegetation at the forest-tundra ecotone. The Holocene 18:323-332

Sugita S (2007a) Theory of quantitative reconstruction of vegetation I: pollen from large sites REVEALS regional vegetation composition. *The Holocene* 17:229-241

Sugita S (2007b) Theory of quantitative reconstruction of vegetation II: all you need is LOVE. The Holocene 17:243-257

Trondman A-K, Gaillard M-J, Sugita S, Mazier F, Fyfe R, Lechterbeck J, Marquer L, Nielsen AB, Twiddle C, Barratt P, Birks HJB, Bjune AE, Caseldine C, David R, Dodson J, Dörfler, W, Fischer E, Giesecke T, Hultberg T, Kangur M, Kuneš P, Latalowa M, Leydet M, Lindbaldh M, Mitchell F, Odgaard B, Peglar SM, Persson T, Rösch M, van der Knaap P, van Geel B, Smith A, Wick L (2015) Pollen-based land-cover reconstructions for the study of past vegetation-climate interactions in NW Europe at 0.2 k, 0.5 k, 3 k and 6 k years before present. Global Change Biology 21:676-697

Tzedakis PC (2007) Seven ambiguities in the Mediterranean palaeoenvironmental narrative. Quaternary Science Reviews 26:2042-2066

von Post L (1946) The prospect for pollen analysis in the study of the earth's climatic history. New Phytologist 45:193-193

Vannière B, Colombaroli D, Roberts N (2010) A fire paradox in ecosystems around the Mediterranean. PAGES News 18:63-65.

Walker D (1970) Direction and rate of change of some British post-glacial hydroseres. In: Walker D, West RG (eds) Studies in the vegetation history of the British Isles. Cambridge University Press, Cambridge, pp. 117-39.

Walker MJC, Berkelhammer M, Björck S, Cwynar LC, Fisher DA, Long AJ (2012) Formal subdivision of the Holocene Series/Epoch: a Discussion Paper by a Working Group of INTIMATE (Integration of ice-core, marine and terrestrial records) and the Subcommission on Quaternary Stratigraphy (International Commission on Stratigraphy). Journal of Quaternary Science 27:649-659

Ward JH (1963) Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58:236-244

Williams JW, Blois JL, Shuman BN (2011) Extrinsic and intrinsic forcing of abrupt ecological change: case studies from the late Quaternary. Journal of Ecology 99:664-677

623 Williams JW, Jackson ST (2007) Novel climates, no analog communities, and ecological
624 surprises. Frontiers in Ecology and the Environment 5:475-482

625 Williams M, Zalasiewicz J, Haff PK, Schwägerl C, Barnosky AD, Ellis EC (2015) The
626 Anthropocene biosphere. The Anthropocene Review 2:196-219

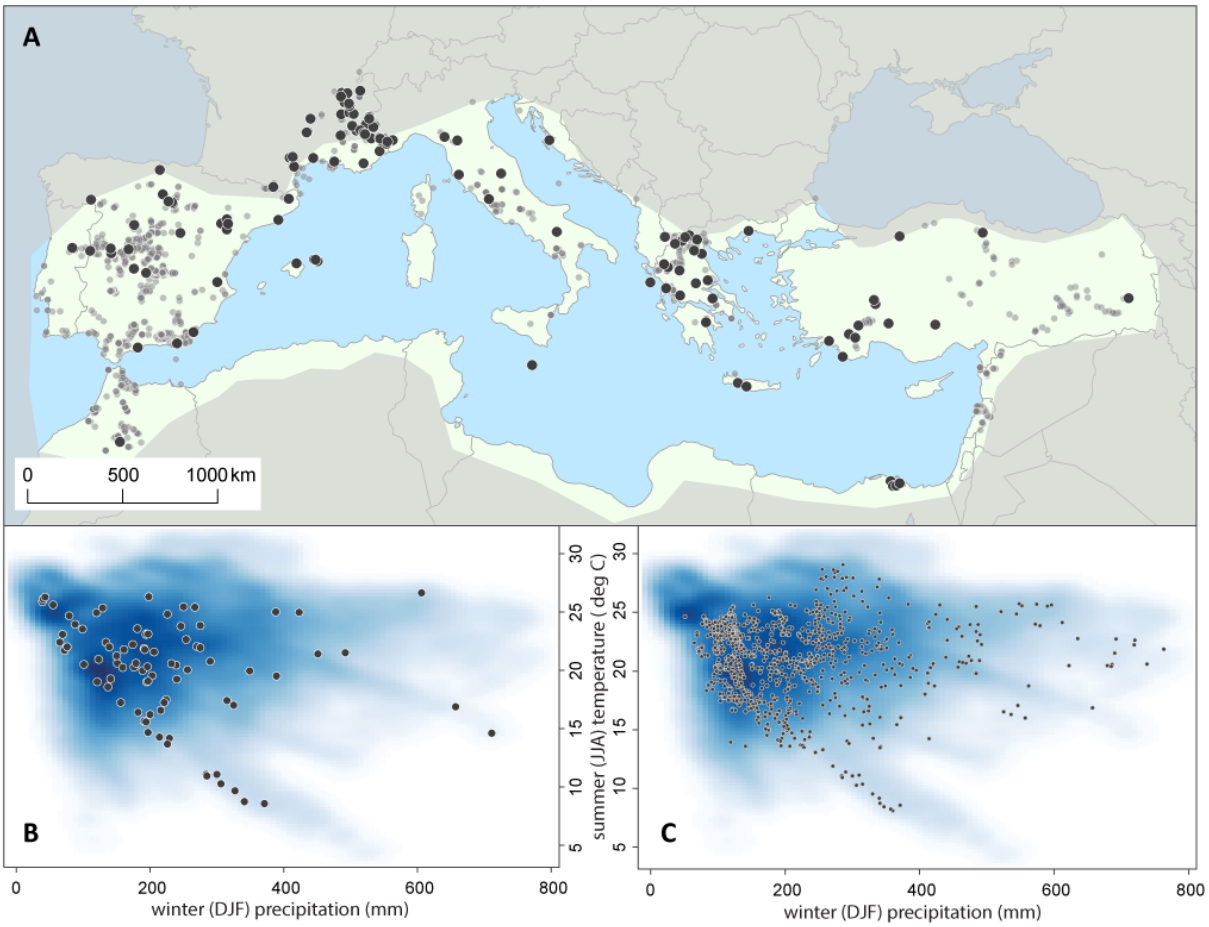
627 Woodbridge J, Fyfe RM, Roberts N (2014) A comparison of remotely-sensed and pollen-
628 based approaches to mapping Europe's land cover. Journal of Biogeography 41:2080-
629 2092

630 Woodbridge J, Roberts N, Fyfe RM (in review) Perpetual change? Holocene vegetation and
631 land cover dynamics in the Mediterranean from pollen data. Journal of Biogeography

632

Uncorrected author's copy. Cite as: Fyfe RM, Woodbridge J and Roberts CN (2017) Trajectories of change in Mediterranean Holocene vegetation through classification of pollen data. *Vegetation History and Archaeobotany* doi:10.1007/s00334-017-0657-4

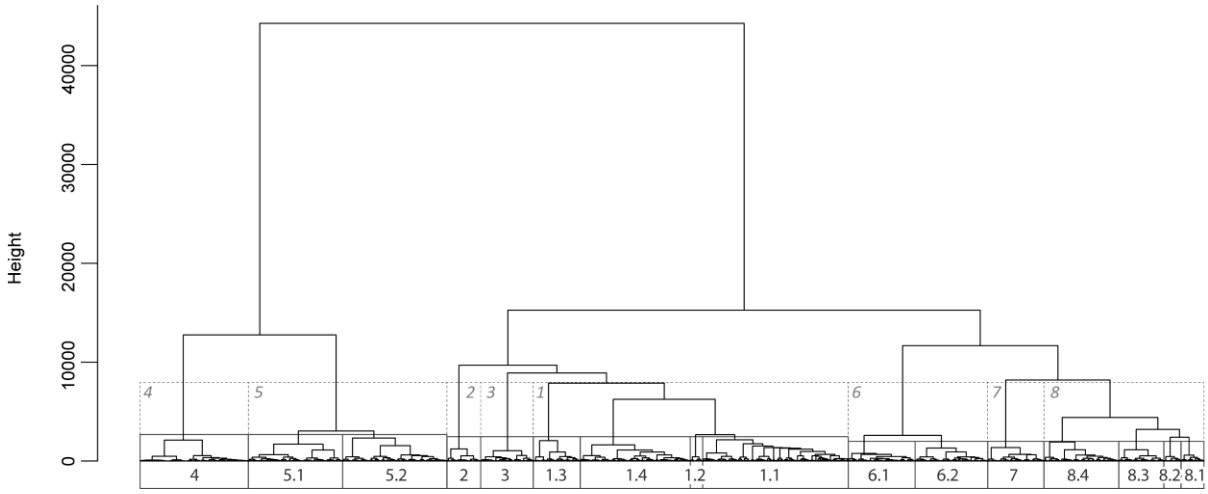
633 Figure 1: Location of sites from the Mediterranean region from the European Pollen Database
634 and European Surface Sample Pollen Database within geographical (A) and climatological
635 space (B: EPD sites; C: EMPD surface samples). In (B) and (C) overall Mediterranean
636 climate is described using a kernel density plot of all grid cells within the Mediterranean
637 region (data taken from Hijmans et al. 2005).
638



639
640

Uncorrected author's copy. Cite as: Fyfe RM, Woodbridge J and Roberts CN (2017) Trajectories of change in Mediterranean Holocene vegetation through classification of pollen data. *Vegetation History and Archaeobotany* doi:10.1007/s00334-017-0657-4

641 Figure 2: Dendrogram derived from hierarchical clustering of the pollen samples using
642 Ward's method. The final 16 derived clusters are indicated with solid boxes; the higher level
643 of aggregation to 8 'meta-clusters' is indicated with dotted boxes.
644

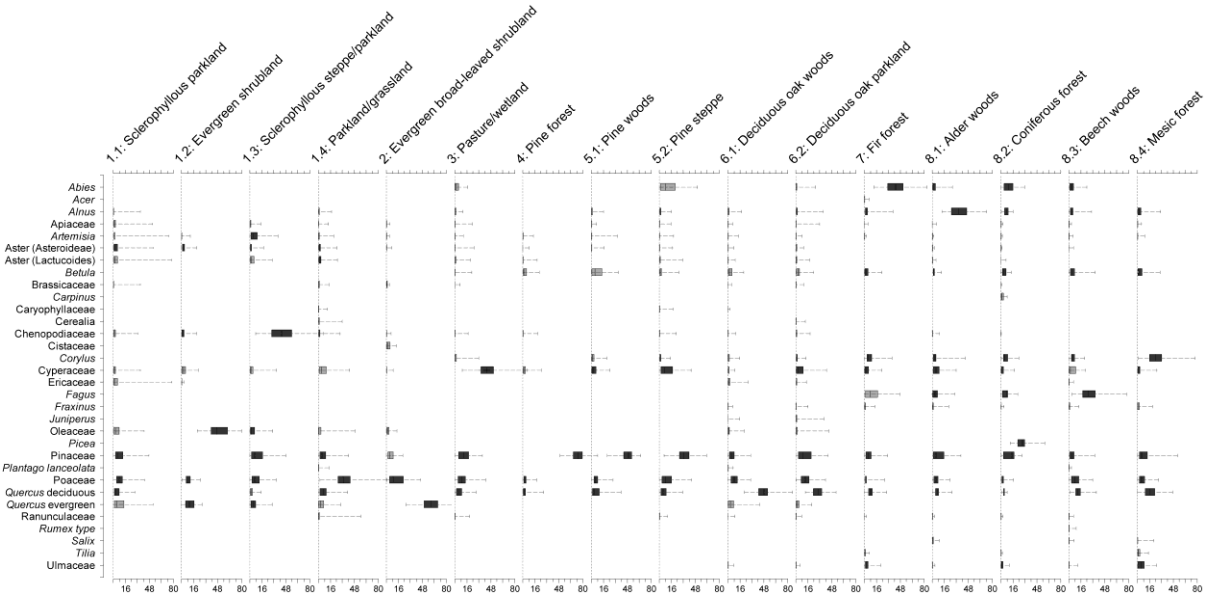


645

646

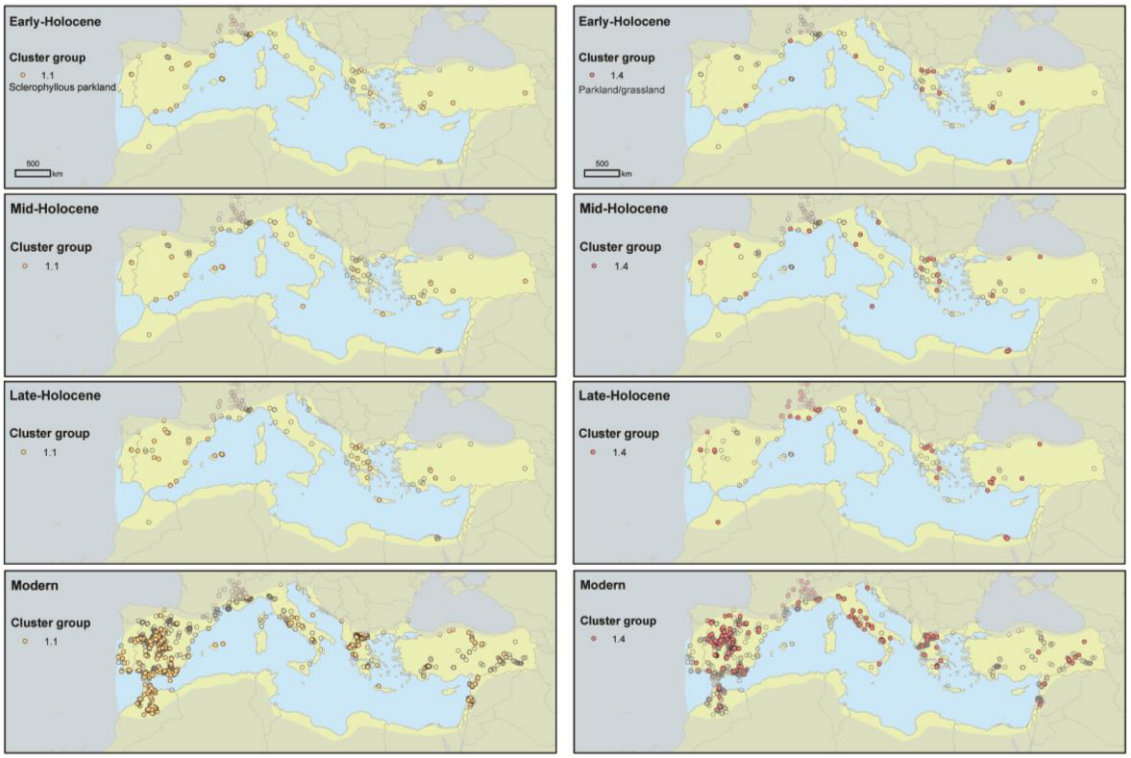
Uncorrected author's copy. Cite as: Fyfe RM, Woodbridge J and Roberts CN (2017) Trajectories of change in Mediterranean Holocene vegetation through classification of pollen data. *Vegetation History and Archaeobotany* doi:10.1007/s00334-017-0657-4

Figure 3: Constant taxa in each of the sixteen cluster groupings derived from Ward's hierarchical clustering method. Only constant taxa (frequency groups V (dark shading) and IV (mid-grey shading)) are shown. Boxes indicate the interquartile range of each taxon within the cluster, with the median, maximum and minimum values also shown.



Uncorrected author's copy. Cite as: Fyfe RM, Woodbridge J and Roberts CN (2017) Trajectories of change in Mediterranean Holocene vegetation through classification of pollen data. *Vegetation History and Archaeobotany* doi:10.1007/s00334-017-0657-4

Figure 4a,b: Spatial and temporal distribution of cluster groups 1.1, 1.4, 5.1 and 6.1. Samples are aggregated into broad time windows to show patterning between the early-, mid- and late-Holocene.



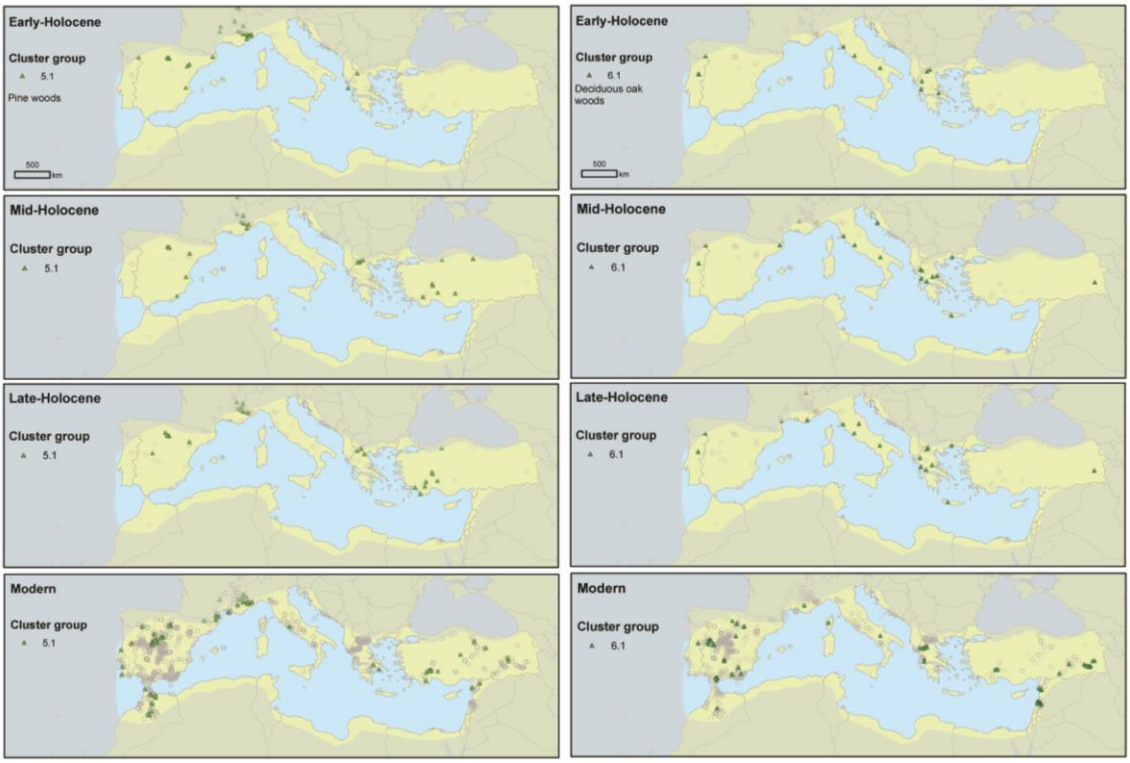
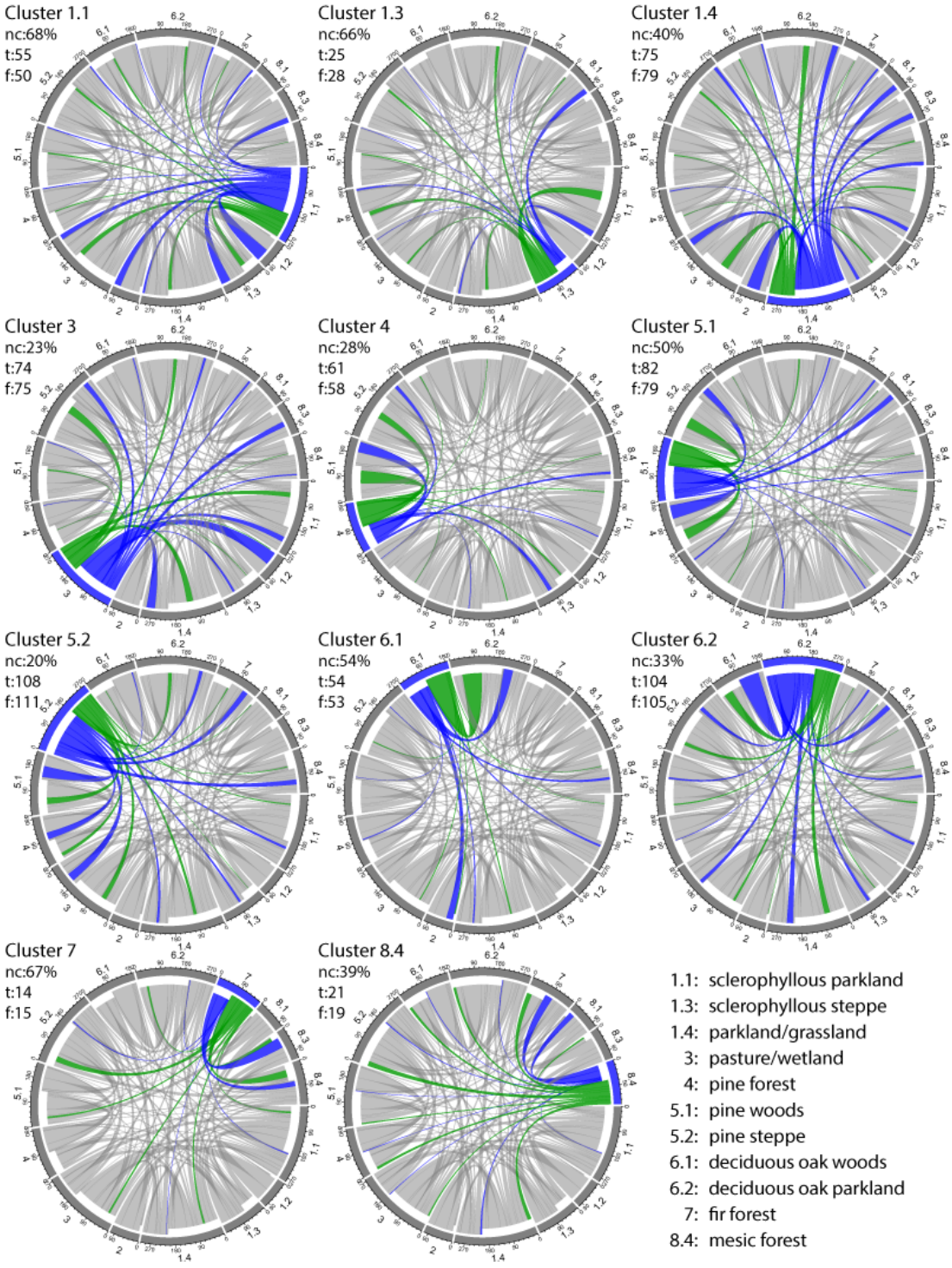


Figure 5: Chord diagrams for each pollen cluster indicating frequency and direction of change for each group. Blue lines flow towards the cluster, indicating the precursor groups. Green lines flow from each cluster, and indicate the groupings to which samples in each site switch. Values of 't' and 'f' on each plot indicate total number of switches to and from each cluster; 'nc' indicates percentage of samples that do not switch across consecutive time windows.



Uncorrected author's copy. Cite as: Fyfe RM, Woodbridge J and Roberts CN (2017) Trajectories of change in Mediterranean Holocene vegetation through classification of pollen data. *Vegetation History and Archaeobotany* doi:10.1007/s00334-017-0657-4

671 Table 1: Descriptions of the cluster groups, including dominant taxa and summaries of the human
672 impact indices used (OJC: Oleaceae, *Juglans*, *Castanea sensu* Mercuri et al., 2013a; API: *Artemisia*,
673 *Centaurea*, Lactuceae [syn. Cichorieae], *Plantago*, cereals, *Urtica*, *Trifolium sensu* Mercuri et al.,
674 2013b) and ranking of clusters by human impact indices.

Cluster	Cluster name (dominant taxa)	%OJC index median (IQR)	OJC rank	%API index median (IQR)	API rank
1.1	Sclerophyllous parkland (Poaceae-Quercus-Oleaceae-herbaceous and sclerophyll taxa)	2.42 (0.56-8.47)	2	7.89 (2.45-22.57)	2
1.2	Evergreen shrubland (Oleaceae-Quercus evergreen-Poaceae-herbaceous and sclerophyllous taxa)	46.78 (39.05-61.8)	1	6.27 (3.03-9.76)	4
1.3	Sclerophyllous steppe/parkland (Chenopodiaceae-Poaceae-Artemisia-Quercus evergreen-Oleaceae)	1.85 (0.45-5.88)	4	8.64 (4.96-16.3)	1
1.4	Parkland/grassland (Poaceae, Quercus deciduous, Asteraceae)	0.97 (0.23-3.44)	5	6.32 (3.4-12.29)	3
2	Evergreen broad-leaved shrubland (Quercus evergreen, Poaceae)	2.14 (0.79-4.41)	3	2.42 (1.44-4.05)	8
3	Pasture/wetland (Cyperaceae, Poaceae, Quercus deciduous)	0.42 (0-2.46)	8	3.26 (0.96-6.96)	6
4	Pine forest (Pinaceae)	0.03 (0-0.73)	13	1.57 (0.79-3.08)	13
5.1	Pine woods (Pinaceae-Poaceae-Quercus deciduous)	0.14 (0-0.68)	9.5	2.15 (0.94-4.56)	9
5.2	Pine steppe (Pinaceae-Poaceae-Quercus deciduous-Artemisia)	0.13 (0-0.96)	11	3.44 (1.53-7.52)	5
6.1	Deciduous oak woods (Quercus-Poaceae-Pinaceae)	0.64 (0.07-2.74)	6	1.62 (0.85-3.07)	12
6.2	Deciduous oak parkland (Quercus-Poaceae-Pinaceae-sclerophyll taxa)	0.43 (6.06-2.07)	7	3.04 (13.13-6.16)	7
7	Fir forest (Abies-Quercus deciduous-Corylus-Pinaceae)	0 (0-0)	15	0.5 (0.17-1.06)	16
8.1	Alder woods (Alnus-Quercus deciduous-Cyperaceae)	0.06 (0-0.55)	12	1.91 (0.79-3.55)	10
8.2	Coniferous forest (Picea-Pinaceae-Abies)	0 (0-0.3)	15	1.2 (0.74-2.04)	14
8.3	Beech woods (Fagus-Quercus deciduous-Poaceae-Corylus)	0.14 (0-0.67)	9.5	1.72 (0.96-3.24)	11
8.4	Mesic forest (Corylus-Quercus deciduous-Poaceae-Pinaceae-Ulmaceae)	0 (0-0)	15	1.05 (0.45-2.86)	15

Uncorrected author's copy. Cite as: Fyfe RM, Woodbridge J and Roberts CN (2017) Trajectories of change in Mediterranean Holocene vegetation through classification of pollen data. *Vegetation History and Archaeobotany* doi:10.1007/s00334-017-0657-4

675 Table 2: Summary of dissimilarity matrix scores for each cluster, and dissimilarity scores showing
676 difference between clusters. Values are based on average Euclidean distance between all samples
677 within, or between, clusters. Higher scores indicate a greater degree of dissimilarity.

	1.1	**1.2**	**1.3**	**1.4**	**2**	**3**	**4**	**5.1**	**5.2**	**6.1**	**6.2**	**7**	**8.1**	**8.2**	**8.3**	**8.4**
1.1	**53.4**															
1.2	62.6	**28.0**														
1.3	63.4	70.8	**38.9**													
1.4	52.7	62.1	62.3	**34.1**												
2	67.4	72.9	78.5	66.5	**26.7**											
3	61.1	70.6	69.7	55.0	78.3	**30.5**										
4	78.2	91.8	84.3	80.0	94.0	79.6	**18.6**									
5.1	60.0	74.7	68.5	60.2	78.0	62.1	35.3	**28.2**								
5.2	53.8	67.8	62.7	50.2	73.3	51.3	51.0	36.9	**34.1**							
6.1	61.2	72.1	72.1	57.3	76.3	66.1	85.6	65.4	59.7	**25.7**						
6.2	52.0	64.3	63.1	45.6	70.4	53.2	72.5	52.3	45.9	35.6	**31.4**					
7	64.2	74.3	73.6	63.5	80.7	66.0	84.6	66.5	55.0	65.0	58.0	**30.8**				
8.1	59.5	70.5	69.3	57.5	76.3	60.7	79.0	61.3	54.5	62.1	52.6	61.5	**32.5**			
8.2	55.2	67.1	65.3	54.8	74.0	59.1	72.7	54.5	47.9	60.3	50.4	51.6	51.2	**23.9**		
8.3	54.8	66.6	65.7	50.7	73.4	57.6	79.0	58.8	51.1	53.1	45.2	53.0	51.7	46.0	**29.9**	
8.4	55.9	68.5	66.9	53.1	75.0	59.4	76.5	57.0	51.4	49.9	44.1	56.9	53.6	49.0	46.3	**35.0**

678

679